THE BERKELEY GROUP

The Berkeley Group
Final Presentation

Nicole Benun, Matt Campbell, Joao Drummond, Achilleas Ghinis, Lauren Hanlon, Emily Rah, Justin Rezende, Taylor Lyberger

TBG Overview

Work

The Berkeley Group provides pro bono consulting services for nonprofit organizations and social enterprises throughout the San Francisco Bay Area.

Mission

TBG strives to provide the highest quality services to our clients in order to maximize their capacity for social impact.

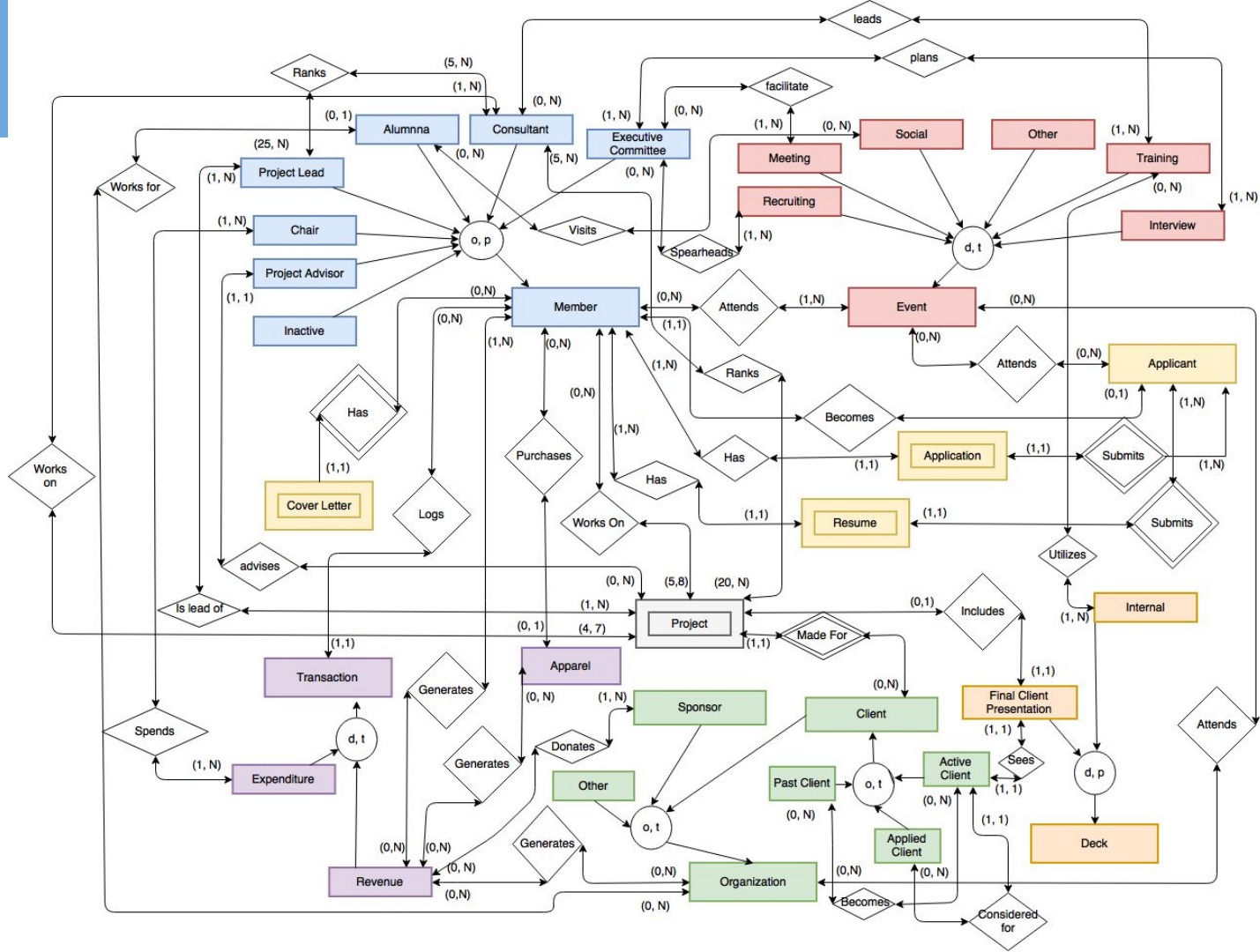
History

TBG was founded in 2003 by four UC Berkeley students. They hoped to provide an opportunity for students to grow professionally and personally, as well as contribute to the growing social sector.

TBG Services

Updated EER

- *Members*
- *Events*
- *Organizations*
- *Decks*
- *Transactions*
- *Applicants*
- *Projects*



Relations: Entities

1. Member(Member_ID, M_fname, M_lname, M_Email, M_Phone, M_Address, Graduation_year, Major(s), Semester_joined, Applicant_ID[6], Gender, Race, Status)

2. Events(Event_ID, Organizer[1], E_date, E_time, E_location, ASUC_sponsored, Est_attendance, Budget, Adv_avenues)

3. Organization(Organization_ID, O_name, O_address, O_phone, O_email)

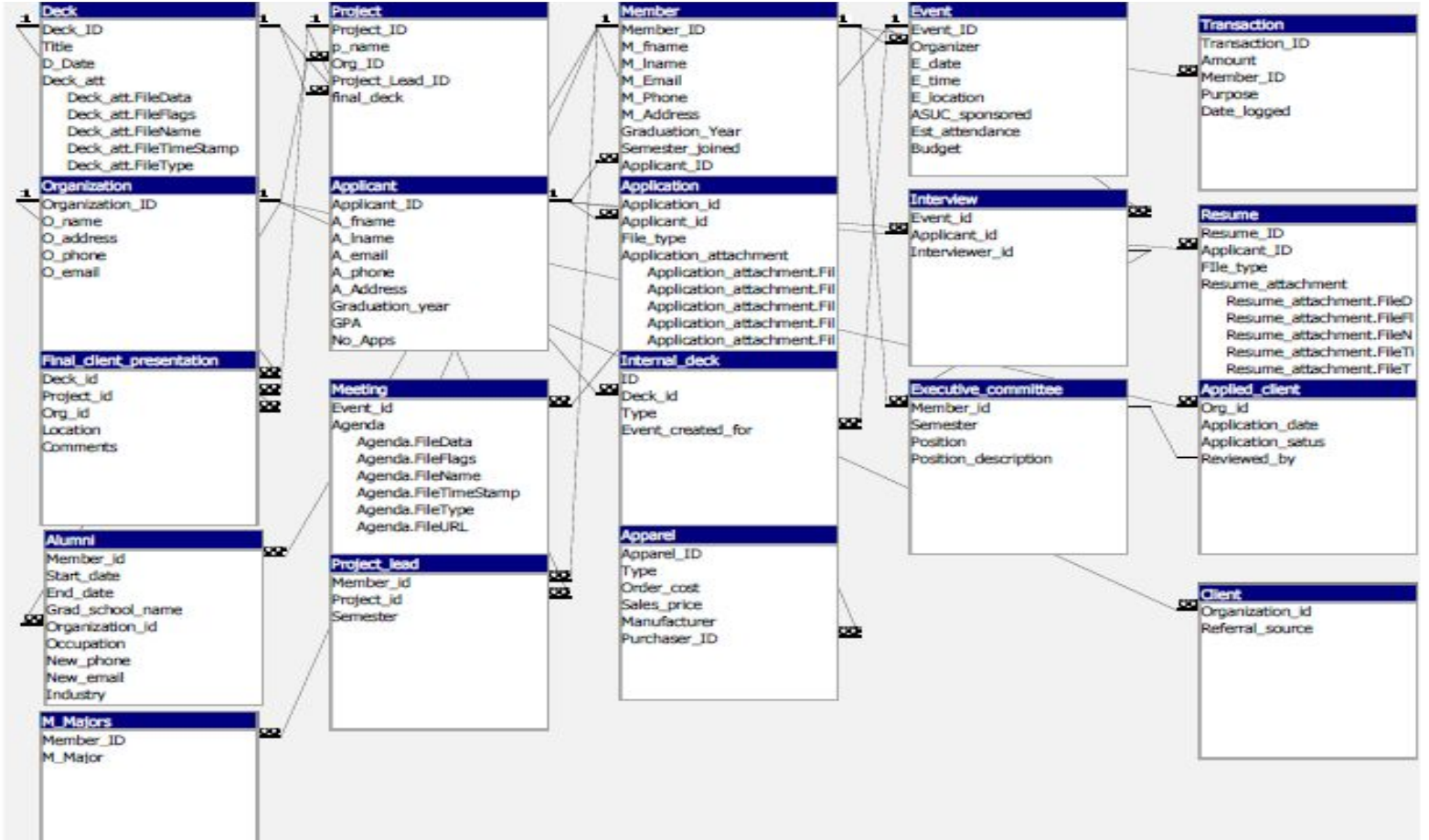
4. Deck(Deck_ID, Title, Author(s), D_Date)

5. Transaction(Transaction_ID, Amount, Member_ID[1], Purpose, Date_logged

6. Applicant(Applicant_ID, A_fname, A_lname, A_Email, A_Phone, A_Address, Graduation_year, Major(s), GPA, No_Apps, Status, Teamwork_Score, Presentation_Score, Quant_Score, Nonprofit_Score, Compatibility_Score)

12. Project(Project_ID,p_name, Organization_ID[3], Project_Lead_ID[1D], project_member(s)[1], final_deck[4], comment(s))

Access Relationships



Normalization Analysis

	Relation	Violation	Fix
2NF	Leads(Member_ID, Event_ID, Date, Training_deck, Semester)	Partial Dependency: Event_ID → Date	Member_Leads(Member_ID, Event_ID) Event_Led(Event_ID, Date, Training_deck, Semester)
3NF	Donates(Organization_ID, Transaction_ID, Type, Amount, Date, Semester)	Non-prime attributes depended on each other: Date → Semester	Donates(Organization_ID, Transaction_ID, Type, Amount, Date) Donate_Sem(Date, Semester)

Relational Schema

Violations: 1NF 2NF 3NF BCNF

MEMBERS

1. **Member**(Member_ID, M_fname, M_lname, M_Email, M_Phone, M_Address, Graduation_year, Semester_joined, Applicant_ID[6], Gender, Race, Status)
   a. **Executive_committee**(Member_ID[1], Semester, Position, Position_description)
   b. **Consultant**(Member_ID[1], Semester, Project_name[12])
   c. **Alumni**(Member_ID[1], Start_date, End_date, Grad_school_name, Organization_name[3], Occupation, New_phone, New_email, Industry)
   d. **Project_lead**(Member_ID[1], Semester, Project_name[12])
   e. **Chair**(Member_ID[1], Semester, Position, Position_description)
   f. **Project_advisor**(Member_ID[1], Semester, Lead_ID[15], Project_ID[12])
   g. **Inactive**(Member_ID[1], Semester_inactive, Number_semesters_inactive, Number_semesters_active)

ORGANIZATIONS

3. **Organization**(Organization_ID, o_name, o_address, o_phone, o_email)
   a. **Client**(Client_Organization_ID[3], Source_Ref)
      i. **Active_client**(Act_Organization_ID[3], active_date, comments(s))
      ii. **Applied_client**(App_Organization_ID[3], application_date, application_status, comment(s), reviewed_by)
      iii. **Past_client**(Organization_ID[3], date_last_active, comment(s), growth_metric)
   b. **Sponsor**(Organization_ID[3], type)
   c. **Other**(Organization_ID[3], comment(s))

EVENTS

2. **Events**(Event_ID, Organizer[1], E_date, E_time, E_location, ASUC_sponsored, Est_attendance, Budget, Advertising_avenues)
   a. **Recruiting**(Event_ID[2], Lead_member[14])
   b. **Interview**(Event_ID[2], Applicant_ID[6], Interview_ID[14])
   c. **Meeting**(Event_ID[2], agenda)
   d. **Social**(Event_ID[2], alumni_invited)
   e. **Training**(Event_ID[2], training_name, presenter[1])
   f. **Other**(Event_ID[2], Other_name)

DECKS

4. **Deck**(Deck_ID, title, date)
   a. **Internal**(Deck_ID[4], type, event_created for[2])
   b. **Final_client_presentation**(Deck_ID[4], Project_ID[12], Organization_ID[3], location, feedback)

TRANSACTIONS

5. **Transaction**(Transaction_ID, amount, Member_ID[1], purpose, date_logged)
   a. **Expenditure**(Transaction_ID[5], date_spent, date_reimbursed)
   b. **Revenue**(Transaction_ID[5], date_given)

Relational Schema

Violations: 1NF 2NF 3NF BCNF

APPLICANTS

6. **Applicant**(Applicant_ID, A_fname, A_lname, A_Email, A_Phone, A_Address, Graduation_year, Major(s), GPA, No_Apps, Status, Teamwork_Score, Presentation_Score, Quant_Score, Nonprofit_Score, Compatibility_Score)
7. **Application**(Application_ID, Applicant_ID[6], File_type, Application_attachment, Source_Ref)
8. **Resume**(Resume_ID, Applicant_ID[6], File_type, Resume_attachment)
9. **Cover_letter**(CL_ID, Applicant_ID[6], File_type, CL_attachment)
10. **Transcript**(Transcript_ID, Applicant_ID[6], File_type, Transcript_attachment)

APPAREL

11. **Apparel**(Apparel_ID, type, order_cost, sale_price, manufacturer, Purchaser_ID[1])

PROJECTS

12. **Project**(Project_ID,p_name, Organization_ID[3], Project_Lead_ID[1D], project_member(s)[1], semester, final_deck[4], comment(s))

MULTIVALUED ATTRIBUTES

31. **M_Majors**(Member_ID[1], M_Major)
32. **Tranining_names**(Member_ID[1], Training_name)
33. **Authors**(Deck_ID[4], Author)
34. **A_Majors**(Applicant_ID[6], Major)
35. **Advertising_avenues**(Event_ID[2], Avenue)
36. **Project_members**(Project_ID[12], project_member)

RELATIONSHIPS

13. **Leads**(Member_ID[1B], Event_ID[2E], Date, Training_deck, Semester)
14. **Plans**(Member_ID[1A], Event_ID[2B], Scheduling_system, Location, Semester)
15. **Facilitates**(Member_ID[1A], Event_ID[2C], Planning_resources, Meeting_deck)
16. **Spearheads**(Member_ID[1A], Event_ID[2A], Recruiting_role, Hours_contributed, Date)
17. **Member_Attends**(Member_ID[1], Event_ID[2], Date, Clock-in, Clock-out)
18. **Applicant_Attends**(Applicant_ID[6], Event_ID[2], Date, Referral_source)
19. **Works_On**(Member_ID[1], Project_ID[12], Start_date, End_date, Semester)
20. **Utilizes**(Event_ID[2E], Deck_ID[4A], Creation_date, Update_date, Semester)
21. **Made_for**(Project_ID[12], Organization_ID[3A], Project_scope, Semester, Contact)
22. **Spends**(Member_ID[1E], Transaction_ID[5A], Item, Price, Quantity, Money_source)
23. **Member_Generates**(Member_ID[1], Transaction_ID[5B], exp_type, date, semester)
24. **Donates**(Organization_ID[3B], Transaction_ID[5], type, amount, date, semester)
25. **Org_Attends**(Organization_ID[3], Event_ID[2], date, representative)
26. **App_Generates**(Transaction_ID[5B], Apparel_ID[11], type, amount)
27. **Org_Generates**(Transaction_ID[5B], Organization_ID[3], amount, type)
28. **Becomes**(App_Organization_ID[3ai], Act_Organization_ID[3ai], sem_Accepted)
29. **Consultant_Rank**(Project_lead[1D], Consultant[1B], Rank, Semester)
30. **Project_Rank**(Consultant[1B], Project_lead[12], Project, Rank, Semester)

Query 1 - Applicant Score Overview

Business Objective

Create a quantitative measure of applicant's abilities to help facilitate discussion during recruitment

Methodology

1. **Holistically score large amount of fake applicants, and individual components of their application**
2. **Perform multilinear regression in STATA or MATLAB on results to determine weights of different factors**
3. **Score Applicants: [GPA, Major Difficulty, Teamwork, Presentation Skills, Quantitative Abilities, Nonprofit Passion, Cultural Fit]**

In Practice

Score applicants throughout the recruitment process, and pull the data during cuts, where scores can be calculated in Excel.

Query 1 - Applicant Score SQL

Applicant_ID	Major
2	5
3	2
4	5
5	5
6	5
7	4

Applicant_ID	A_fname	A_lname	A_e	A_ph	A_Addı	Graduati	GPA	No_App	Status	Teamwork_S	Presentation_	Quant_Score	Nonprofit_Sc	Compatibility
2	Justin	Rezende	jrezen	301-65	2395 Pied	12/16/2017	3.5	4	Active	9	10	10	5	4
3	Matt	Robertson	mrobe	310-12	1234 Colli	5/7/2017	3.73	5	Active	10	9	7	8	7
4	Nicole	Benun	nbenu	310-99	2400 Fult	5/12/2018	3.8	2	Active	9	8	9	6	6
5	Achilleas	Ghinis	aghini	510-98	2532 Colli	5/18/2017	2.7	1	Active	4	7	5	8	3
6	Joao	Drummond	joaoga	510508	2301 Dura	5/18/2017	3.3	2	Active	6	6	8	4	7
7	Taylor	Lyberger	tlyberg	310850	2709 Chai	5/18/2017	3.9	3	Active	7	4	9	8	8

```
SELECT Applicant.A_fname, Applicant.A_lname, Applicant.GPA,
Applicant.Teamwork_Score, Applicant.Presentation_Score,
Applicant.Quant_Score, Applicant.Nonprofit_Score,
Applicant.Compatibility_Score, A_Majors.Major
FROM Applicant INNER JOIN A_Majors ON Applicant.[Applicant_ID]
= A_Majors.[Applicant_ID];
```

A_fname	A_lname	GPA	Teamwork_S	Presentation_	Quant_Score	Nonprofit_Sc	Compatibility	Major
Justin	Rezende	3.5	9	10	10	5	4	5
Matt	Robertson	3.73	10	9	7	8	7	2
Nicole	Benun	3.8	9	8	9	6	6	5
Achilleas	Ghinis	2.7	4	7	5	8	3	5
Joao	Drummond	3.3	6	6	8	4	7	5
Taylor	Lyberger	3.9	7	4	9	8	8	4

Query 1 - Applicant Score In Practice

	A	B	C	D	E	F	G	H	I
1	A_fname	A_lname	GPA	Teamwork_Score	resentation_Scor	Quant_Score	Nonprofit_Score	ompatibility_Scor	Major
2	Justin	Rezende	3.5	9	10	10	5	4	5
3	Matt	Robertson	3.73	10	9	7	8	7	2
4	Nicole	Benun	3.8	9	8	9	6	6	5
5	Achilleas	Ghinis	2.7	4	7	5	8	3	5
6	Joao	Drummond	3.3	6	6	8	4	7	5
7	Taylor	Lyberger	3.9	7	4	9	8	8	4
8									

Conversion, Weighting, Ranking

- GPA
- Teamwork
- Presentation
- Quantitative
- Nonprofit Passion
- Cultural Fit
- Major Difficulty

	A	B	J
1	A_fname	A_lname	Score
2	Matt	Robertson	402
3	Nicole	Benun	378
4	Justin	Rezende	365
5	Taylor	Lyberger	356
6	Achilleas	Ghinis	299
7	Joao	Drummond	297

Query 2 – Marketing Strategy Success Overview

Business Objective

Understand which recruitment efforts are most efficiently increasing gross applications

Methodology

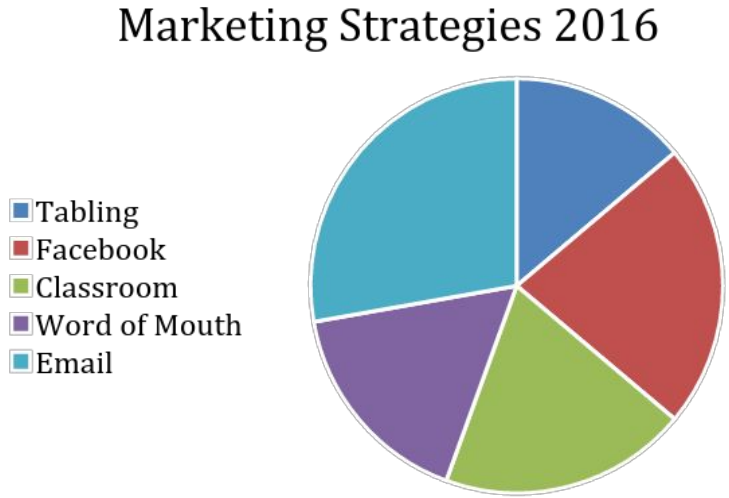
1. **Track how applicants heard about TBG**
2. **Determine hours of work spent on each type of recruitment**
3. **Export applicant data and compare with hours worked**

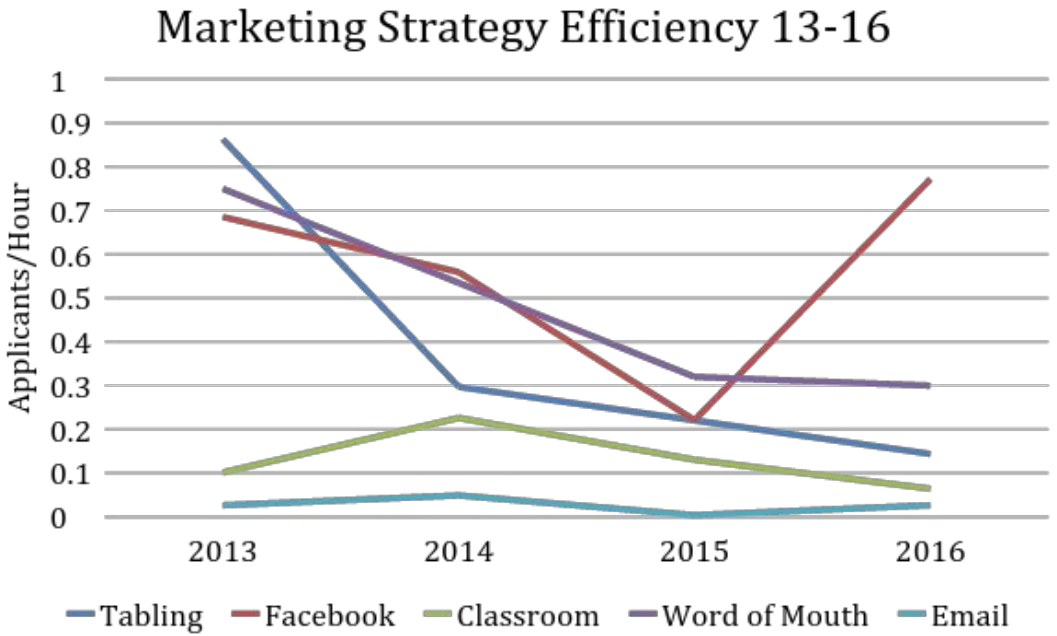
In Practice

Guide TBG towards focusing on most efficient recruitment practices

Query 2 – Marketing Strategy Success

```
SELECT Application.[Referral_Source], Application.[Semester],
Count(Application.Referral_Source) AS CountOfReferral_Source
FROM Application
GROUP BY Application.[Referral_Source],
Application.[Semester];
```





Query 3 – Alumni Matching Overview

Business Objective

Connect current undergraduate members to other members and alumni, who would be most relevant and helpful as mentors

Methodology

1. **User inputs either an industry or company**
2. **Pull members and alumni matching the specified company or industry**
3. **Rank by projects worked on together, then shared events, then graduation year (relatability to current member)**

In Practice

Streamline identifying potential alumni for application help and career development

Query 3 – Alumni Matching SQL

```
@M_fname = 'Justin'
@M_lname = 'Rezende'
@Organization_name = 'Bain'
SQL>  SELECT Grad_school_name, Organization_name, Occupation, Industry,
New_email Member_ID AS 'Alumni_ID'
FROM Alumni
WHERE Member_ID IN (SELECT Member_ID FROM Alumni WHERE Organization_name =
@Organization_name  FROM Alumni);
SQL >  CREATE TABLE Events_attended AS
SELECT a.Member_ID, DISTINCT a.Event_ID
FROM Attends a, Member m
WHERE a.Member_ID == (SELECT Member_ID FROM Member WHERE M_fname == @M_fname and
M_lname == @M_lname  FROM Member);
SQL > SELECT a.Member_ID
FROM Alumni a
WHERE a.Member_ID IN (SELECT Member_ID FROM Events_attended);
```

Query 4 – Project Matching Overview

Business Objective

Automate and optimize the assignment of consultants to project teams based on consultant and project lead preferences

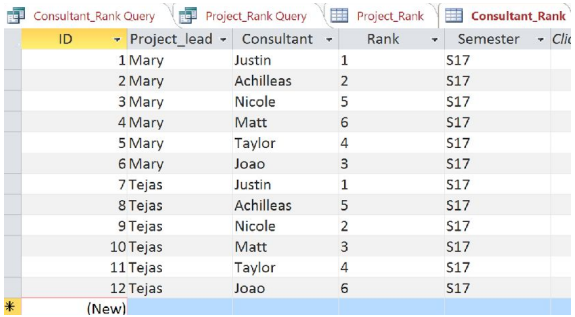
Methodology

1. **Consultants rank the projects in a semester and submit apps**
2. **Project leaders rank consultants based on fit for project and abilities**
3. **Query preferences and export as .csv to R to run Happy Marriage algorithm**

In Practice

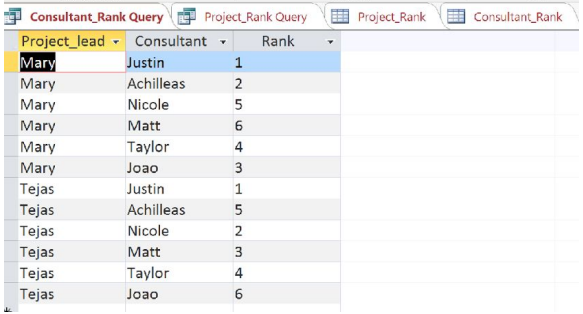
Fast and effective initial creation of project teams

Query 4 – Project Matching SQL

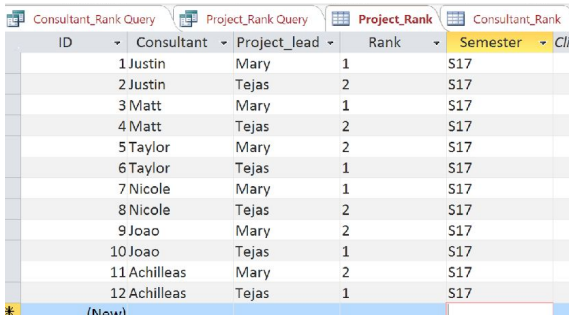


ID	Project_lead	Consultant	Rank	Semester
1	Mary	Justin	1	S17
2	Mary	Achilleas	2	S17
3	Mary	Nicole	5	S17
4	Mary	Matt	6	S17
5	Mary	Taylor	4	S17
6	Mary	Joao	3	S17
7	Tejas	Justin	1	S17
8	Tejas	Achilleas	5	S17
9	Tejas	Nicole	2	S17
10	Tejas	Matt	3	S17
11	Tejas	Taylor	4	S17
12	Tejas	Joao	6	S17

```
SELECT Consultant_Rank.Project_lead,
Consultant_Rank.Consultant,
Consultant_Rank.Rank
FROM Consultant_Rank;
```

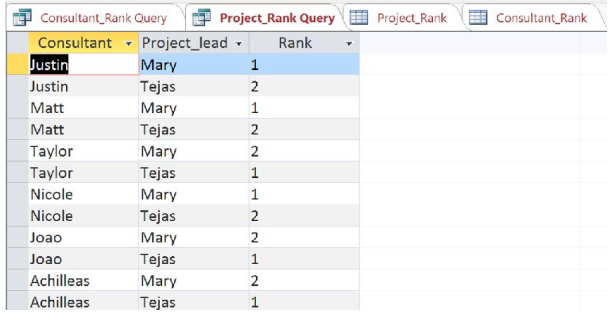


Project_lead	Consultant	Rank
Mary	Justin	1
Mary	Achilleas	2
Mary	Nicole	5
Mary	Matt	6
Mary	Taylor	4
Mary	Joao	3
Tejas	Justin	1
Tejas	Achilleas	5
Tejas	Nicole	2
Tejas	Matt	3
Tejas	Taylor	4
Tejas	Joao	6



ID	Consultant	Project_lead	Rank	Semester
1	Justin	Mary	1	S17
2	Justin	Tejas	2	S17
3	Matt	Mary	1	S17
4	Matt	Tejas	2	S17
5	Taylor	Mary	2	S17
6	Taylor	Tejas	1	S17
7	Nicole	Mary	1	S17
8	Nicole	Tejas	2	S17
9	Joao	Mary	2	S17
10	Joao	Tejas	1	S17
11	Achilleas	Mary	2	S17
12	Achilleas	Tejas	1	S17

```
SELECT Project_Rank.Consultant,
Project_Rank.Project_lead,
Project_Rank.Rank
FROM Project_Rank;
```



Consultant	Project_lead	Rank
Justin	Mary	1
Justin	Tejas	2
Matt	Mary	1
Matt	Tejas	2
Taylor	Mary	2
Taylor	Tejas	1
Nicole	Mary	1
Nicole	Tejas	2
Joao	Mary	2
Joao	Tejas	1
Achilleas	Mary	2
Achilleas	Tejas	1

Query 4 – Project Matching In Practice

Import to R

```
##    Project_lead Consultant  Rank
##           (chr)      (chr) (dbl)
## 1         Mary     Justin     1
## 2         Mary  Achilleas     2
## 3         Mary     Nicole     5
## 4         Mary      Matt      6
## 5         Mary     Taylor     4
## 6         Mary      Joao      3
## 7        Tejas     Justin     1
## 8        Tejas  Achilleas     5
## 9        Tejas     Nicole     2
## 10       Tejas      Matt      3
## 11       Tejas     Taylor     4
## 12       Tejas      Joao      6
```

```
##         (chr)        (chr) (dbl)
## 1      Justin         Mary     1
## 2      Justin        Tejas     2
## 3        Matt         Mary     1
## 4        Matt        Tejas     2
## 5      Taylor         Mary     2
## 6      Taylor        Tejas     1
## 7      Nicole         Mary     1
## 8      Nicole        Tejas     2
## 9        Joao         Mary     2
## 10       Joao        Tejas     1
## 11  Achilleas         Mary     2
## 12  Achilleas        Tejas     1
```

Wrangle Data

```
consultant_matches_final
```

```
##   Consultant Project_Lead
## 1  Achilleas         Mary
## 2       Joao         Mary
## 3     Justin        Tejas
## 4       Matt        Tejas
## 5     Nicole        Tejas
## 6     Taylor         Mary
```

```
project_lead_matches_final
```

```
##   Project_Lead Consultant
## 1         Mary     Taylor
## 2         Mary       Joao
## 3         Mary  Achilleas
## 4        Tejas       Matt
## 5        Tejas     Nicole
## 6        Tejas     Justin
```

Create “Happy” Teams

```
## Source: local data frame [6 x 2]
##
##    Mary Tejas
##   (dbl) (dbl)
## 1     2     5
## 2     3     6
## 3     1     1
## 4     6     3
## 5     5     2
## 6     4     4
```

```
pref_cons
```

```
## Source: local data frame [2 x 6]
##
##   Achilleas  Joao Justin  Matt Nicole Taylor
##       (dbl) (dbl)  (dbl) (dbl)  (dbl)  (dbl)
## 1         2     2      1     1      1      2
## 2         1     1      2     2      2      1
```

Query 5 – Budget Calculation and Forecasting Overview

Business Objective

Help TBG track spending over course of semester, and roughly forecast spending based on previouss semester

Methodology

1. **Pull data from past 3 years of equivalent semesters spending**
2. **Export to made to input Excel sheet**
3. **Forecast based on amount spent so far in the semester end of the semester expenditures**

In Practice

Throughout the semester tracking of budget usage and effect on club's finances over time

Query 5 – Budget Calculation and Forecasting SQL

```
SELECT Transaction.[Amount], Month(Transaction.[Date_logged]),
Year(Transaction.[Date_logged])
FROM Transaction
WHERE Year(Transaction.[Date_logged]) >
Year(Transaction.[Date_logged]) - 3;
```

Thank you! *Questions?*

Relational Schema

1. Member(Member_ID, M_fname, M_lname, M_Email, M_Phone, M_Address, Graduation_year, Major(s), Semester_joined, Applicant_ID[6], Gender, Race)
1a. Executive_committee(Member_ID[1], Semester, Position, Position_description)
1b. Consultant(Member_ID[1], Semester, Project_name[12], Training_name(s)[2e])
1c. Alumni(Member_ID[1], Start_date, End_date, Grad_school_name, Organization_name[3], Occupation, New_phone, New_email, Industry)
1d. Project_lead(Member_ID[1], Semester, Project_name[12])
1e. Chair(Member_ID[1], Semester, Position, Position_description)
1f. Project_advisor(Member_ID[1], Semester, Lead_ID[1D], Project_name[12])
1g. Inactive(Member_ID[1], Semester_inactive, Number_semesters_inactive, Number_semesters_active)

2. Events(Event_ID, E_date, E_time, E_location, ASUC_sponsored, est_attendance, budget
2a. Recruiting(Event_ID[2], Lead_member[1a])
2b. Interview(Event_ID[2], Applicant_ID[6], Interviewer_ID[1a])
2c. Meeting(Event_ID[2], agenda)
2d. Social(Event_ID[2], alumni_invited)
2e. Training(Event_ID[2], training_name, Presenter[1])
2f. Other(Event_ID[2])

Relational Schema

3. Organization(Organization_ID,o_name,o_address,o_phone,o_email)
3a. Client(Organization_ID[3], Referral_Source)
3ai. Active_client(Organization_ID[3],active_date, comments(s))
3aii. Applied_client(Organization_ID[3],application_date,application_status,comment(s), reviewed_by[1a])
3aiii. Past_client(Organization_ID[3],date_last_active,comment(s), growth_metric)
3b. Sponsor(Organization_ID[3],type)
3c. Other(Organization_ID[3], comment(s))

4. Deck(Deck_ID,title,author(s),date)
4a. Internal_deck(Deck_ID[4], type, event_created_for[2])
4b. Final_client_presentation(Deck_ID[4],Project_ID[**],Organization_ID[3],location,comment(s))

5. Transaction(Transaction_ID, amount, Member_ID[1], use, date_logged)
5a. Expenditure(Transaction_ID[5], date_spent, date_reimbursed)
5b. Revenue(Transaction_ID[5], date_given)

Relational Schema

6. Applicant(Applicant_ID, A_fname, A_lname, A_Email, A_Phone, A_Address, Graduation_year, Major(s), GPA, No_Apps, Status, Teamwork_Score, Presentation_Score, Quant_Score, Nonprofit_Score, Compatibility_Score)

7. Application(Application_ID, Applicant_ID[6], File_type, Application_attachment, Referral_Source, Semester)

8. Resume(Resume_ID, Applicant_ID[6], File_type, Resume_attachment)

9. Cover_letter(CL_ID, Applicant_ID[6], File_type, CL_attachment)

10. Transcript(Transcript_ID, Applicant_ID[6], File_type, Transcript_attachment)

11. Apparel(Apparel_ID, type, order_cost, sale_price, manufacturer, Purchaser_ID[1])

12. Project(Project_ID,p_name, Organization_ID[3], Project_Lead_ID[1D], project_member(s)[1], final_deck[4], comment(s))

13. Project_App_Rank(PL_ID[1D], Applying_ID[1B], Rank)
13a. Project_Lead_Score(PL_ID[1D], Applying_ID[1B], Score)

Relational Schema

14. **Leads**(Member_ID[1B], Event_ID[2E])
15. **Plans**(Member_ID[1A], Event_ID[2B])
16. **Facilitates(**Member_ID[1A], Event_ID[2C])
17. **Spearheads**(Member_ID[1A], Event_ID[2A])
18. **Attends**(Member_ID[1], Event_ID[2])
19. **Works_On**(Member_ID[1], Project_ID[12])
20. **Utilizes**(Event_ID[2E], Deck_ID[4A])
21. **Made_for**(Project_ID[12], Organization_ID[3A]
22. **Spends**(Member_ID[1E], Transaction_ID[5A])
23. **Generates(**Member_ID[1], Transaction_ID[5B])
24. **Donates**(Organization_ID[3B], Transaction_ID[5])
25. **Org_Attends**(Organization_ID[3], Event_ID[2])
26. **App_Generates**(Transaction_ID[5B], Apparel_ID[11])
27. **Org_Generates**(Transaction_ID[5B], Organization_ID[3])
28. **Becomes**(Organization_ID[3])

Relations: Multivalued attributes

29. **M_Majors**(Member_ID[1], M_Major)
30. **Tranining_names**(Member_ID[1],Training_name)
31. **Authors(**Deck_ID[4], Author)
32. **A_Majors**(Applicant_ID[6], Major)
33. **A_App_IDs**(Applicant_ID[6], Application_ID)
34. **Project_members**(Project_ID[12], project_member)

Sum the value assignments to retrieve the “distance” between two members